

Sistem Pendukung Keputusan untuk Menentukan Jurusan Siswa Baru di SMK Negeri 1 Kragilan Menggunakan Metode Naive Bayes

Dian Tri Budianto¹, Agus Irawan²

Program Study Strata Satu Teknik Informatika
Universitas Serang Raya

¹dyanzzz@gmail.com, ²agus.irawan@unsera.ac.id

Abstrak – SMKN 1 Kragilan merupakan Sekolah Menengah Kejuruan yang memberikan pendidikan kepada siswa – siswinya dalam hal kemampuan khusus sesuai jurusan yang diminatinya, dan bertujuan meningkatkan kualitas siswa – siswi sebelum masuk kedalam dunia kerja. Salah satu hal penting yang dapat menentukan kualitas siswa di dalam dunia kerja yaitu dapat masuk kedalam jurusan yang sesuai dan diminati oleh siswa dan siswi tersebut. Dalam pemilihan jurusan dilakukan pada awal siswa masuk sekolah ketika mendaftarkan diri di SMKN 1 Kragilan, penjurusan didasari dengan nilai UN, yang sudah ditentukan sebelumnya untuk persyaratan. Diharapkan dengan dukungan system yang dibangun ini dapat menyelesaikan permasalahan dalam pengambilan keputusan untuk menentukan jurusan siswa baru, agar siswa tidak salah dalam memilih jurusan yang sesuai dengan minat dan bakatnya.

Kata Kunci : SMKN 1 Kragilan, System Pendukung Keputusan, Naive Bayes, System Penentuan Jurusan

1. PENDAHULUAN

Lembaga pendidikan setiap tahunnya terus berusaha meningkatkan mutu pendidikan khususnya Sekolah Menengah Kejuruan (SMK) dengan harapan lulusannya dapat memiliki ketrampilan dan keahlian lebih dibandingkan sekolah sederajat, hal tersebut dilakukan demi meningkatkan kualitas lulusan sehingga siap memasuki dunia kerja. Antusias siswa lulusan SMP untuk masuk SMK cukup besar, tetapi kebanyakan mereka kurang matang untuk memilih jurusan yang ada sesuai kemampuannya, akibatnya cukup banyak siswa baru yang gagal di tengah jalan ketika mereka sudah diterima di SMK, banyak juga kasus siswa yang merasa tidak cocok dengan jurusan yang dipilihnya ketika ia telah memperoleh pelajaran disekolah siswa tersebut tidak merasa nyaman dengan pelajaran yang ditempuh. [1]

2. DASAR TEORI

2.1. DATA MINING

Data mining adalah penambangan atau penemuan informasi baru dengan mencari pola atau aturan tertentu dari sejumlah data yang sangat besar. Data mining juga disebut sebagai serangkaian proses untuk menggali nilai tambah berupa pengetahuan yang selama ini tidak diketahui secara manual dari suatu kumpulan data. [2]

2.2. METODE KLASIFIKASI

Klasifikasi adalah proses untuk menemukan model atau fungsi yang menjelaskan atau membedakan konsep atau kelas data, dengan tujuan untuk dapat memperkirakan kelas dari suatu objek yang labelnya tidak diketahui. Dalam mencapai tujuan tersebut, proses klasifikasi membentuk suatu model yang mampu membedakan data kedalam kelas-kelas yang berbeda berdasarkan aturan atau fungsi tertentu. Model itu sendiri bisa berupa aturan “jika-maka”, berupa pohon keputusan, atau formula matematis. [3]

2.3. TEOREMA BAYES

Dalam teori probabilitas dan statistika, teorema Bayes adalah sebuah teorema dengan dua penafsiran berbeda. Dalam penafsiran Bayes, teorema ini menyatakan seberapa jauh derajat kepercayaan subjektif harus berubah secara rasional ketika ada petunjuk baru. Dalam penafsiran frekuentis teorema ini menjelaskan representasi invers probabilitas dua kejadian. Teorema ini merupakan dasar dari statistika Bayes dan memiliki penerapan dalam sains, rekayasa, ilmu ekonomi (terutama ilmu ekonomi mikro), teori permainan, kedokteran dan hukum. Penerapan Teorema Bayes untuk memperbarui kepercayaan dinamakan inferens Bayes. [4]

2.4. NAIVE BAYES

Algoritma *Naive Bayes* merupakan salah satu algoritma yang terdapat pada teknik klasifikasi. *Naive Bayes* merupakan pengklasifikasian dengan metode probabilitas dan statistik yang dikemukakan oleh ilmuwan Inggris *Thomas Bayes*, yaitu memprediksi peluang di masa depan berdasarkan pengalaman dimasa sebelumnya sehingga dikenal sebagai *Teorema Bayes*. Teorema tersebut dikombinasikan dengan *Naive* dimana diasumsikan kondisi antar atribut saling bebas. Klasifikasi *Naive Bayes* diasumsikan bahwa ada atau tidak ciri tertentu dari sebuah kelas tidak ada hubungannya dengan ciri dari kelas lainnya. Dasar dari metode *Naive Bayes* adalah *teorema bayes* yang menyatakan bahwa, jika X merupakan sampel data kelas (label) yang tidak diketahui, dan H adalah hipotesa dimana X merupakan data kelas (label) C , dan $P(H)$ adalah peluang dari hipotesa H , kemudian $P(X)$ dinyatakan sebagai peluang kejadian X (data sample) yang diamati, maka $P(X|H)$ adalah peluang data sample X , yang diasumsikan bahwa hipotesa H bernilai benar (valid). [4]

3. Design Penelitian

3.1. Metodologi

Jenis Penelitian yang dilaksanakan adalah jenis penelitian informasi kuantitatif yaitu pengambilan data yang telah dikumpulkan oleh SMKN 1 Kragilan.

3.2. Sampling

NO	JURUSAN	NILAI UN			
		B. IND	B. ING	MTK	IPA
1	AK	8.6	6.6	5	5.25
2	AK	8.8	6.2	4.5	4.5
3	AK	8.4	7.5	7.7	8.2
4	AK	6.6	4.4	8.5	7
5	AK	7.4	4.8	4.5	4
6	AK	5	3.6	4	3.25
7	AK	6	3.8	2.5	7.25
8	AK	5	3.2	2.5	2.75
9	AK	8	4.2	3.5	4
10	AK	5.4	4.2	2.5	3
11	AK	3.8	6.8	7.25	6.75
12	AK	5.4	5	7.25	4.25
13	AK	8.4	7.5	7.8	8.2
14	AK	5	3.2	3.25	3
15	AK	6.4	4	4	3
16	AK	5.6	3.8	2	4
17	AK	5.4	1.8	2.75	6.75
18	APK	4.6	6.4	3.75	6.5
19	APK	7.8	5.8	5.5	5.25
20	APK	7.6	5.2	3.75	2
21	APK	4	5.8	3.5	2.75
22	APK	8.2	4.4	3	3
23	APK	7.2	4	2.75	4.25
24	APK	8	5	3	3
25	APK	6.2	3.8	3.25	3.5
26	APK	3.2	6.2	2	5.5
27	APK	7.6	2.6	3.25	5
28	APK	6	3.8	3.25	5.25
29	APK	5.8	4	2.25	2.5
30	APK	8.2	6.4	6.5	5.25
31	APK	7.8	4	3.25	4.25
32	APK	7.4	3	2.5	4.5
33	APK	7	3.8	3.5	3.75
34	APK	5.2	4	3.25	3.75
35	APK	5.2	3.6	7.25	7
36	APK	6	3.2	1.25	3.75
37	APK	8.2	4	2.5	4
38	APK	4.6	3.8	4.25	4.5
39	APK	5.4	3.4	3.75	7.25
40	APK	6.2	3.6	2.75	3.25
41	APK	7.4	3.6	2.75	2.5
42	APK	6.4	4.2	2.25	3.25
43	APK	6.4	6	7.25	2
44	APK	3.4	3.6	8	2.75
45	MESIN	6	6	3.75	7.25
46	MESIN	4.6	4.6	3.75	7.75
47	MESIN	4	4	3.25	4
48	MESIN	4.8	4.8	3	2.25
49	MESIN	4.6	4.6	4.75	3
50	MESIN	5.2	5.2	7.75	4.25
51	MESIN	7	7	2.5	2.25
52	MESIN	2.6	2.6	7.25	4
53	MESIN	2	2	3.25	4.5
54	MESIN	3.4	3.4	2.25	6.5
55	MESIN	3.6	3.6	2	7
56	MESIN	6.2	6.2	6	3.25
57	MESIN	5.8	5.8	4.5	7.4
58	MESIN	7.4	7.4	6.5	2.75
59	MESIN	3.2	3.2	3	7
60	MESIN	7.6	7.6	5.25	5.1
61	MESIN	4.4	4.4	7.75	6.5
62	MESIN	4	4	3	3.25
63	MESIN	5.4	5.4	3.5	7.25
64	MESIN	3.4	3.4	4	3.25
65	MESIN	3.8	3.8	3.25	7
66	MESIN	2.6	2.6	3.75	4.25
67	MESIN	4.6	4.6	4.75	3.25
68	MESIN	8.7	8.7	7.8	7.75
69	MESIN	2.8	2.8	2.25	7
70	MESIN	3.8	3.8	3.75	3.75
71	MESIN	5.4	5.4	3.5	4.75
72	MESIN	6.4	6.4	6.75	6
73	MESIN	2.8	2.8	3.5	3.5
74	MESIN	5.2	5.2	6.75	3.75
75	MESIN	3.2	3.2	3.25	3.25
76	MESIN	4.8	4.8	1.75	6.75
77	MESIN	3.2	3.2	6.75	6.75
78	MESIN	6.4	6.4	7.25	8
79	OTO	5.6	3.4	4	3
80	OTO	6.8	3.4	3.25	3.75
81	OTO	8.2	6.4	4.2	4.75
82	OTO	6.2	4.4	3.75	3
83	OTO	7	5	4.75	4.75
84	OTO	7.8	3.4	3.75	4
85	OTO	6.2	3.6	2.5	2.5
86	OTO	5.4	4.4	3	4.75
87	OTO	6.4	2.6	3.75	4.25
88	OTO	5.2	4.2	7	3
89	OTO	8.2	7	7	8
90	OTO	6	3.4	3	4.5
91	OTO	6.8	4.6	5	6
92	OTO	5.4	3.2	3.5	3.75
93	OTO	6.8	3	3	3.5
94	OTO	5.4	5.4	2	3
95	OTO	7.2	4.6	2.5	3.75
96	OTO	6	3.8	1.75	3.25
97	OTO	6.4	4.2	4.5	3.75
98	OTO	7.2	4.2	3.25	4.25
99	OTO	5.2	3	4	2.25
100	OTO	6	3.6	3.75	3
101	OTO	7.4	4.4	3.25	3.5
102	OTO	8	6	2.75	3
103	OTO	8.4	5.7	8	8.5
104	OTO	8	7.4	7.25	7.25
105	OTO	3.2	5.4	7.75	6.75
106	OTO	7.2	3.2	3.25	2.5
107	OTO	4.6	4.2	1.75	3.5

NO	JURUSAN	NILAI UN				NO	JURUSAN	NILAI UN			
		B. IND	B. ING	MTK	IPA			B. IND	B. ING	MTK	IPA
108	OTO	5.6	3.8	3.5	4.5	163	TKJ	7	4.2	3	3
109	OTO	8	7.2	3	3.25	164	TKJ	3.2	6	7.25	3.5
110	OTO	4.4	5.8	5.25	4	165	TKJ	7	5	4	3.5
111	OTO	7.6	6.8	6.25	7	166	TKJ	7.2	5.6	5.75	3.75
112	OTO	7.8	4	4	4.2	167	TKJ	7.2	4.2	4	3.5
113	OTO	7.8	6	2	5.5	168	TKJ	5	4.2	5.25	4.5
114	OTO	6.6	3.8	2.75	3.75	169	TKJ	3.2	7.2	3.5	3.5
115	OTO	6	4.2	2	3.75	170	TKJ	6.8	6.8	7.25	6.25
116	OTO	3	5.4	6	7.25	171	TKJ	8.6	3.6	3.25	4
117	OTO	6	4.8	6.75	6.75	172	TKJ	6.2	3.2	3.25	3.75
118	OTO	6.4	3.8	2.5	5.5	173	TKJ	6.4	3.8	4.5	4.25
119	OTO	5.2	3	8.5	2.75	174	TKJ	4.8	3.2	7	6
120	OTO	7.6	3.8	2.75	3.5	175	TKJ	8.8	5.2	4.25	6.5
121	OTO	7.2	3.8	1.75	5	176	TKJ	3.4	7.4	5	8
122	OTO	8	4	2.5	4.75	177	TKJ	7.6	2	2.75	4.5
123	OTO	4.8	2.8	2.5	4.25	178	TKJ	7.2	3.8	2	3.5
124	OTO	4.6	3.6	4.5	3.25	179	TKJ	7.2	3.2	8.25	3
125	OTO	5.4	2.4	6	7.5	180	TKJ	7.8	2.2	3	2.25
126	OTO	5.8	8.6	8.25	3	181	TKJ	6.8	4.2	3.75	3.25
127	OTO	4.2	3	7	5	182	TKJ	6.2	4.4	4.25	4
128	OTO	7	3.2	4.25	5.25	183	TKJ	4.2	7.2	2.75	3.25
129	OTO	4.4	4.4	2.5	6.75	184	TKJ	6.8	4.8	3	3.5
130	OTO	6.6	4.4	3	4.75	185	TKJ	5.4	5.2	3	8.25
131	OTO	6.6	3.6	3.5	3.5						
132	OTO	5.2	4.4	3.5	3.5						
133	TKJ	8.6	6.2	2.5	3.75						
134	TKJ	9	5.4	7.25	3.25						
135	TKJ	7.8	7.4	3.5	5.25						
136	TKJ	7.6	4.2	3	4.75						
137	TKJ	7.4	6.4	5.75	7.5						
138	TKJ	8	4.4	7.25	3.75						
139	TKJ	7.8	5.8	3.75	3.5						
140	TKJ	9	7.6	4.5	4.75						
141	TKJ	8.6	5	7.75	5						
142	TKJ	7.8	3.6	3.25	4						
143	TKJ	8	4.6	3.75	4						
144	TKJ	8.2	5.6	2	4.5						
145	TKJ	7.8	6.6	5.7	6						
146	TKJ	7.2	6	3	4.5						
147	TKJ	6.4	7.5	8.2	8.7						
148	TKJ	7.8	4.8	5.5	4.5						
149	TKJ	8.4	7.1	7.7	8.6						
150	TKJ	8.2	4.6	3.5	5						
151	TKJ	7	3	4	3.25						
152	TKJ	6.8	4.2	2.25	4						
153	TKJ	7	4.2	4.75	4.25						
154	TKJ	6.8	4.4	4.25	6						
155	TKJ	3.2	5.8	2.75	6.75						
156	TKJ	3.2	3.4	6.75	2.75						
157	TKJ	6	9	9	6.5						
158	TKJ	7	3.2	2	5.25						
159	TKJ	6.4	4.8	3.25	4						
160	TKJ	6	3.8	2.75	2.5						
161	TKJ	6.5	8.1	7.9	7.6						
162	TKJ	6.2	4.6	3.2	3.5						

3.3. Proses Naive Bayes

Jika diketahui ada sebuah data baru yang belum memiliki klas sebagai berikut :

X = (B.Indonesia=5.00, B.Ingggris=7.00, Matematika=8.00, IPA=6.00)

Ubah menjadi Indikator yaitu Rendah<7, Sedang=7-7.9 dan Tinggi>7.9

B.Ind=5.00 : Rendah

B.Ing=7.00 : Sedang

Matematika=8.00 : Tinggi

IPA=6.00 : Rendah

Tentukan klas yang cocok untuk data di atas:

Jawab:

Langkah 1 : Hitung nilai $P(X_k | C_i)$ untuk setiap klas i

$P(B.Ind = \text{"Rendah"} \cap \text{Jurusan} = \text{"AKN"}) = 11$;

$P(\text{Jurusan} = \text{"AKN"}) = 17$

Dengan menggunakan persamaan :

$P(X|H) = \frac{P(X \cap H)}{P(H)}$, maka diperoleh

$$\begin{aligned} P(B.Ind = \text{Rendah} | \text{Jurusan} = \text{"AKN"}) \\ = \frac{P(B.Ind = \text{Rendah} \cap \text{Jurusan} = \text{"AKN"})}{P(\text{Jurusan} = \text{"AKN"})} = \frac{11}{17} \\ = 0.647 \end{aligned}$$

dengan cara yang sama, maka nilai $P(X_k | C_i)$ untuk klas yang lain diperoleh sebagai berikut :

$P(B.Ind = \text{"Rendah"} | \text{Jurusan} = \text{"APK"}) = 15/27 = 0.556$

$P(B.Ind = \text{"Rendah"} | \text{Jurusan} = \text{"MSN"}) = 30/34 = 0.882$

$P(B.Ind = \text{"Rendah"} | \text{Jurusan} = \text{"OTO"}) = 35/54 = 0.648$

$P(B.Ind = \text{"Rendah"} | \text{Jurusan} = \text{"TKJ"}) = 23/53 = 0.434$

$P(B.Ingg = \text{"Sedang"} | \text{Jurusan} = \text{"AKN"}) = 2/17 = 0.118$

$P(B.Ingg = \text{"Sedang"} | \text{Jurusan} = \text{"APK"}) = 0/27 = 0.1$

$$P(B.Ingg="Sedang" | Jurusan="MSN")=3/34=0.088$$

$$P(B.Ingg="Sedang" | Jurusan="OTO")=3/54=0.056$$

$$P(B.Ingg="Sedang" | Jurusan="TKJ")=7/53=0.132$$

$$P(MTK="Tinggi" | Jurusan="AKN")=1/17=0.059$$

$$P(MTK="Tinggi" | Jurusan="APK")=1/27=0.037$$

$$P(MTK="Tinggi" | Jurusan="MSN")=0/34=0.1$$

$$P(MTK="Tinggi" | Jurusan="OTO")=3/54=0.056$$

$$P(MTK="Tinggi" | Jurusan="TKJ")=3/53=0.057$$

$$P(IPA="Rendah" | Jurusan="AKN")=13/17=0.765$$

$$P(IPA="Rendah" | Jurusan="APK")=25/27=0.926$$

$$P(IPA="Rendah" | Jurusan="MSN")=24/34=0.706$$

$$P(IPA="Rendah" | Jurusan="OTO")=48/54=0.889$$

$$P(IPA="Rendah" | Jurusan="TKJ")=47/53=0.887$$

Langkah 2 : Hitunglah nilai $P(X | C_i)$ untuk setiap klas menggunakan persamaan (5)

$$P(X|C_i) = \prod_{k=1}^n P(X_k|C_i), \text{ sehingga di peroleh :}$$

$$P(X | Jurusan="AKN") = 0.647 \times 0.118 \times 0.059 \times 0.765 = 0.0034242885$$

$$P(X | Jurusan="APK") = 0.556 \times 0.1 \times 0.037 \times 0.926 = 0.0019051974$$

$$P(X | Jurusan="MSN") = 0.882 \times 0.088 \times 0.1 \times 0.706 = 0.0054956239$$

$$P(X | Jurusan="OTO") = 0.648 \times 0.056 \times 0.056 \times 0.889 = 0.0017781842$$

$$P(X | Jurusan="TKJ") = 0.434 \times 0.132 \times 0.057 \times 0.887 = 0.0028770109$$

Langkah 3 : Hitunglah nilai klas (label) dari data sample tersebut, gunakan persamaan (6)

$$P(X|C_i) * P(C_i), \text{ sehingga diperoleh :}$$

$$P(X|Jurusan="AKN") * P(Jurusan="AKN") = 0.0034242885 * 0.09189 = \mathbf{0.000314664}$$

$$P(X|Jurusan="APK") * P(Jurusan="APK") = 0.0019051974 * 0.14595 = \mathbf{0.000278056}$$

$$P(X|Jurusan="MSN") * P(Jurusan="MSN") = 0.0054956239 * 0.18378 = \mathbf{0.001010007}$$

$$P(X|Jurusan="OTO") * P(Jurusan="OTO") = 0.0017781842 * 0.29189 = \mathbf{0.000519038}$$

$$P(X|Jurusan="TKJ") * P(Jurusan="TKJ") = 0.0028770109 * 0.28649 = \mathbf{0.000824225}$$

Dengan demikian X memiliki klas "**Jurusan = MSN(Mesin)**" karena memiliki nilai maksimum untuk perhitungan klas (label) nya.

4. Uji Coba System

Sistem dibuat dengan menggunakan php hypertext procesor, berikut Forminputnya:

Gambar 1. Formulir Input Biodata

Gambar 2. Formulir Input Persyaratan dan Nilai UN

Gambar 3. Hasil Klasifikasi dari metode Naive Bayes

5. Kesimpulan dan Saran

5.1. Kesimpulan

Berdasarkan proses pembuatan sistem pendukung keputusan untuk menentukan jurusan, maka proses perancangan *system* yang telah dirancang sebelumnya, diantaranya penulisan analisa data masukan, keluaran dan data proses, perancangan DAD, transformasi ERD ke LRS, pembuatan database struktur table, dan pembuatan layout halaman aplikasi. Tujuan penerapan hasil rancangan agar maksud dan tujuan pembuatan website sistem pendukung keputusan ini dapat tercapai. Setelah penerapan dilakukan maka pada tahap pengujian terhadap pembuatan website *system* pendukung keputusan yang telah dibangun, dilakukan guna mengetahui apakah maksud dan tujuan yang ingin dicapai telah terpenuhi, sehingga dapat ditarik kesimpulan :

1. Dapat memudahkan panitia penerimaan siswa baru untuk memberikan saran kepada siswa dalam mengambil keputusan jurusan yang sesuai dengan minat dan bakat siswa.
2. Dapat mempermudah siswa dalam menentukan pemilihan jurusan ketika siswa tersebut tidak dapat masuk kedalam jurusan yang diinginkan karena persaingan yang begitu erat.

3. Memudahkan pengolahan data pendaftaran sebagai arsip milik sekolah.
4. Memudahkan dan mempercepat proses penerimaan siswa baru.

5.2. Saran

Dari kesimpulan – kesimpulan diatas, maka penyusun mengemukakan saran yaitu:

1. Diharapkan program aplikasi ini dapat dikembangkan kembali agar aplikasi ini digunakan secara lebih baik serta optimal.
2. Diharapkan pada setiap siswa ataupun panitia penerimaan siswa baru dapat menentukan jurusan berdasarkan jurusan yang memang benar-benar sesuai dengan minat dan bakat siswa.

DAFTAR PUSTAKA

1. [Ilyas, Asmidir. 2011]. Padang, Meningkatkan Mutu Pendidikan Melalui Kegiatan Diagnosis Kesulitan Mengajar.
2. [Giarratano, Joseph C. 2005.] *Expert Systems : Principles and Programming*, Thomson Course Technology
3. [Bustami.] Penerapan Algoritma *Naive Bayes* Untuk Mengklasifikasi Data Nasabah Asuransi
4. [TB. Ai Munandar. 2012.] Modul Kecerdasan Buatan Teorema Bayes (AI-6). Universitas Serang Raya: Tidak diterbitkan.